Alternative Image Sensors

redefining the pixel

By Reg Miles

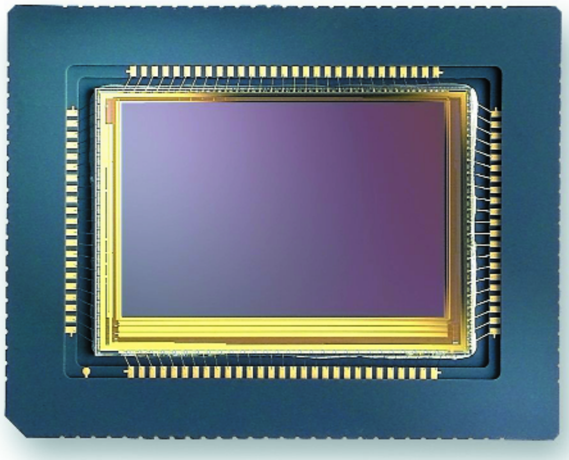
regmiles@tesco.net

Although the vast majority of image sensors in use are of the Charge Coupled Device (CCD) variety (particularly in consumer products), and those that are not are of the Complementary Metal Oxide Semiconductor (CMOS) type; other types are being used in a variety of applications or being developed. A number of these have interesting properties that would suit them to a wider market. And there are also variations on the CCD to adapt it to particular roles.

However, it is worth noting that they all share the same underlying principle of operation. The sensitive area is divided into pixels; where impinging photons can produce electron-hole pairs in the silicon. Either the holes or, more usually, the electrons are then gathered into a potential well while the others are drained into the substrate. The charge in the well is then moved out, amplified, processed and, with the charges from all the other pixels, forms the picture.

Linear sensor

One variation on the common area array sensor is the **linear sensor** — which, as the name implies, is a strip of photosensitive pixels. This type is used in professional digital cameras and scanners. It works by exposing a strip at a time (see **Figure 1**). In each position, the charge is built up (integrated) in the pixels, corresponding to the brightness of the light, the charges are then read out, and the sensor moves to the next position. In practice this is a virtually continuous motion. The advantage of it is that it can cover a relatively large area with far fewer pixels than an area array; the obvious disadvantage is that it is restricted to static subjects. If it is a monochrome shot then only one pass of the sensor is required; if it is a colour shot then it requires three passes, through red, green and blue (RGB) filters in turn.



The alternative is to have a **tri-linear sensor** comprised of three strips of pixels filtered for RGB — thus requiring only one pass.

Linear sensors can provide very high resolution images because of the relatively large area covered. But the resolution of area sensors can also be increased by movement. This can be done by moving either the sensor or the optical image by fractional amounts that will shift it one pixel at a time. Thus, a single, filtered sensor can be shifted three times so that each red, green and blue filtered pixel will be sequentially exposed in the same position — effectively tripling the resolution. The resolution can be increased still further if a monochrome sensor is used and the movements are reduced so that the actual photosensitive area within each pixel is exposed to a different part of the image — gradually filling in the gaps in the sensitivity. Many sensors have pixels that are not uniformly sensitive to light: this ratio of sensitive to non-sensitive is known as the **fill factor**, and can vary from 10% to 100%. If a colour image is required then all the movements will have to be repeated three times.

A variation on the linear sensor is a type used for line scan imaging of, normally, industrial processes. This is the exact opposite of the moving sensor and static image — here, the sensor is static and the image is moving (objects passing on a conveyer belt, for example). To achieve maximum sensitivity, **Time Delay and Integration** (TDI) imaging is normally employed. This utilises more sensor area than a linear array; and shift registers are used to integrate and transport charge in the direction of image motion. Obviously, this has to be done at high speed. For colour, the red, green and blue filtered sections are exposed sequentially.

In an **area array**, high speed is achieved by having multiple outputs that each receive charge from a small area of the sensor, giving frame rates from 1000 to perhaps 5000 per second, depending on the total number of pixels. If 'windowing' is used, with only a portion of the total area being read out, then the frame rate can be 50,000, or more. These techniques are used in conjunction with an electronic shutter, which freezes the image by first dumping the charge and then integrating it for just the required, very brief, period.

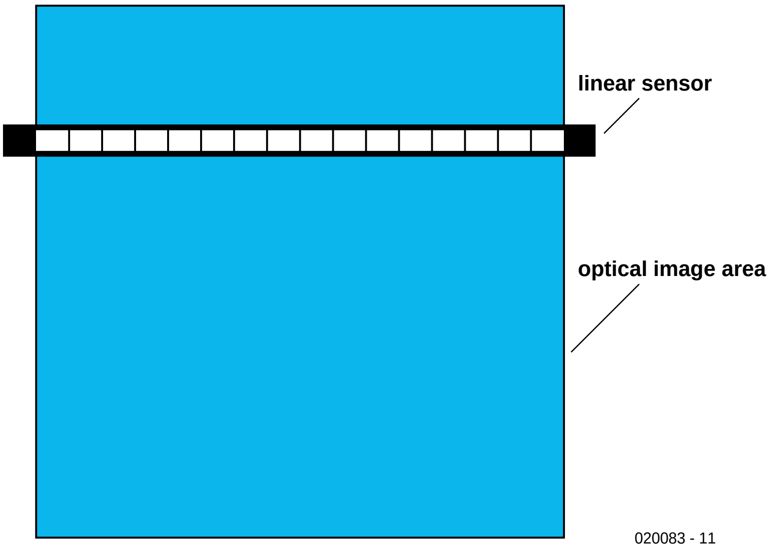


Figure 1. Linear optical sensor principle.

Although high speed sensors are normally CCD types, Stanford University has developed a CMOS **Digital Pixel Sensor** (DPS) which can achieve 10,000 full frames per second. This is made possible by having an 8-bit AD converter in each pixel, which eliminates the analogue bottlenecks. An acceptable pixel size based on DPS requires the use of a 0.18-micron CMOS process; which has produced its own problems due to reduced supply voltages and increased leakage currents. The 9.4-micron square pixels in the 352?288 array use a **photogate detector** (basically a MOS capacitor exposed to light), rather than the more usual p-n junction photodiode detector. Each pixel incorporates 37 transistors, so, not surprisingly, the fill factor is only 15%.

Focal plane processing is also being applied to those sensor applications that normally require an external computer for post-processing of the output. Such things as robotic vision, target acquisition and tracking, car navigation, etc, that rely on edge detection, motion detection, and such like. With each pixel effectively being a tiny computer, the sensor becomes a massively parallel computer, and the need for post processing is either considerably reduced or completely eliminated. The only problem with these 'smart' sensors is, as noted above, the much reduced fill factor.

Advances in focal plane processing

Researchers at CalTech have designed a **focal plane processing** system in which the resolution is enhanced through continuous vibration of the CMOS sensor — filling in the gaps left by the non-sensitive areas of the pixels. Each pixel operates independently, gathering visual information as it scans the scene; and each is read out independently, after in-pixel processing. Having the pixels operating independently eliminates fixed pattern noise because their outputs are not compared. Two means of producing the vibrations have been designed. In one, a 45-degree mirror is used to reflect the image to the lens: the mirror is mounted on a motor shaft slanted at less than 1 degree from the perpendicular so that the image wobbles in a regular circular motion. In the other, the image is irregularly vibrated by the lens being suspended on four springs: the vibrations are either induced naturally — by driving over rough terrain, etc., or by a piezoelectric device.

The Charge Injection Device

One type of sensor that has a high fill factor is the **Charge Injection Device** (CID). Like the CCD it has been under development for a long time: the concept having been invented by researchers working for the General Electric Company, who announced their first CID camera in 1972 (the same year that researchers at Bell Laboratories announced

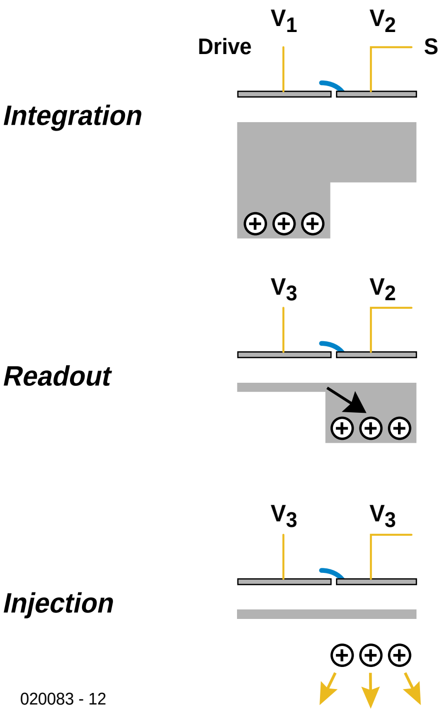


Figure 2. CID pixel dynamics (courtesy Thermo CIDTEC).

CID Pixel Dynamics

- One or more electrodes inverted

$$|V_1| > |V_2| > |V_3| = 0$$

- Change potential on one electrode
- Sense charge transfer (signal) on other electrode. Charge remains within pixel.
- Readout is non-destructive if V_I is restored to Drive electrode
- Injection occurs when both electrode potentials are collapsed ($|V_3| = 0$). Injected charge diffuses and recombines in underlying collector.

the first CCD camera).

The company that is now most associated with the CID technology, Thermo CIDTEC, was established by a management buyout from GEC in 1987.

The CID is more like a CMOS than a CCD, with pixels being individually addressed via row and column electrodes rather than the charges being moved en masse via vertical shift registers. Unlike either, however, each pixel contains two overlapping photogates. Another difference is that the charge is not transferred from the pixel during readout, and so further charge integration can occur — which is particularly useful for adaptive exposure control (selected pixels or groups of pixels can be read several times for real time exposure monitoring during lengthy integration periods). When one photogate is selected by the drive (row) electrode its charge is transferred to the other photogate, and a displacement current proportional to the charge is read by the sense (column) electrode (see **Figure 2**). This is then amplified, converted to a voltage and fed out. The pixel is only emptied when both drive and sense electrodes are switched to ground and the charge is 'injected' into an underlying epitaxial layer. This can be at normal field or frame rate, or at any other rate.

The CID has a number of advantages over other imaging devices. The spilling over of excessive charge from a pixel into neighbouring pixels, causing 'blooming', is minimised because the excess is drawn into the underlying epitaxial layer and, with no shift registers, there are no indirect routes to spread the charge overload. The pixel structure is contiguous. Pixels can hold a lot of electrons (this is called the 'well capacity'). Plus, the device has a broad spectral response due to the absence of colour absorbing shift registers and the minimal layers used to fabricate pixels. On the down side, noise is introduced by connecting the small capacitance of the pixel to the relatively large capacitance of the sense line.

The next step: TFA

Another way of achieving a contiguous pixel structure is by vertical integration of the photodiode on top of the pixel transistors, as used by Silicon Vision for their **Thin Film on ASIC** (TFA) technology. This combines the excellent optical properties of hydrogenated amorphous silicon (a-Si:H) with the advantages of crystalline silicon (x-Si) for integrated circuits, to achieve intelligent image sensors. The thin film of amorphous silicon is deposited on the crystalline silicon Application Specific Integrated Circuit (ASIC) in a Plasma Enhanced Chemical Vapour Deposition process; with anything from four to seven layers being successively deposited, in a thickness of less than one micron. **Figure 3** shows the basic layer sequence of a TFA. The insulation layer is patterned in order to provide the contact holes for the pixels between the detector and the circuitry; with the pixel area being defined by its rear electrode which is a patterned metal layer; the top layer is a transparent conductive oxide that acts as a common electrode for all the pixels.

TFA uses row and column addressing of the pixels; and therefore a variety of readouts are possible, including random access when desired. Pixels may also be connected in ways that enable them to communicate with each other — for such things as compression or pattern recognition. Another advantage

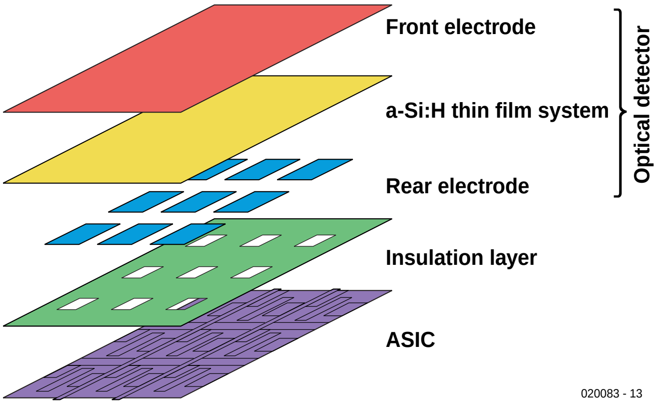


Figure 3. Basic layer sequence of a Thick Film on ASIC. (courtesy Silicon Vision).

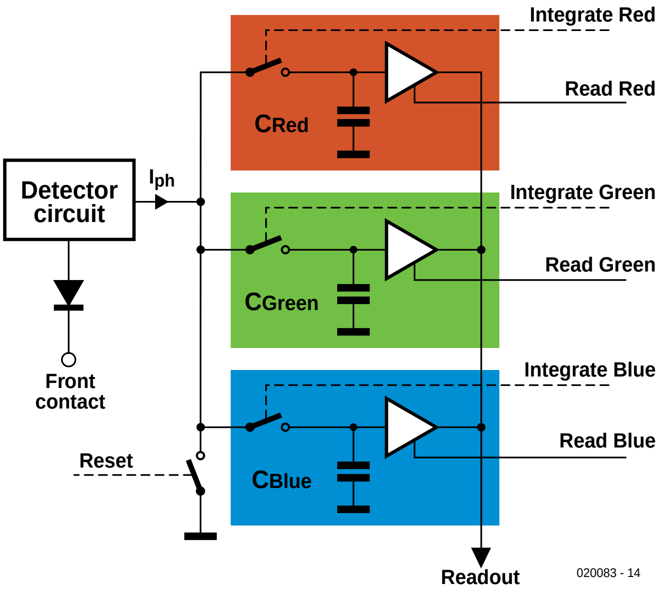


Figure 4. The COSIMA element employs sequential inputting of the colours into their respective integration circuits (courtesy Silicon Vision).

is that the detector and the ASIC can be optimised as if they were separate. Also, the fabrication process is cheaper than that of a CMOS sensor — which is itself cheaper than a CCD.

Silicon Vision has developed three types of image sensor from the basic TFA technology. They are discussed below.

COSIMA, LARS and HIRISE

COSIMA, the Colour Sensor with Integrated Memory Array is, as the name implies, a colour chip; but where it differs from others is that the red, green and blue components are all handled within each individual pixel. Colour recognition is sequential, and is based on the wavelength: blue light is absorbed near the surface of the silicon, green light penetrates deeper, and red deeper still — the actual penetration depths being about 50 nm, 200 nm, and 500 nm to 1 micron. Detection of the three colours is achieved by applying different bias voltages. During the integration phase, the three charges are sequentially fed into their respective colour integration circuits and stored within each pixel (see **Figure 4**). For readout the integrated colour voltages are simultaneously applied to the column output line.

The second type of sensor is the **Lokal-AutoadaptiveR Sensor** (LARS — the company is German). In this, the pixels adapt themselves to the local illumination; and the result is a dynamic range of 120 dB (a normal CCD or CMOS has around 80 dB). Thus suiting it to industrial and automotive vision systems. It can achieve this because it divides the full dynamic range into two signals, each of moderate dynamic range. The integration voltage in the pixel is compared at clocked intervals to a fixed reference voltage, which is slightly below half the saturation value; if it has exceeded that threshold, integration is stopped; if it has not, the integration time is doubled (**Figure 5a** is a diagram of a LARS pixel). Every time the comparison is made a voltage ramp is increased by one step, marking the duration of the integration. When integration is stopped this is sampled and stored in the pixel as timestamp information (**Figure 5b** is a timing diagram of a pixel). When all the pixels have stopped integration, the two voltages – integrated intensity information and timestamp information – are then read out from each; and the original photocurrent is reconstructed by subsequent processing from a combination of the two signals. Colour can be achieved by sequential exposures through red, green and blue filters. The LARS auto-adaptive function can al also be switched off to let the device be used normally.

Thirdly, the high-resolution sensor, **HIRISE**. There are three versions of this. The HIRISE VGA with 640?480 pixels, of 7.4 microns square. The HIRISE HDTV is a 16:9 format sensor with 1920x1040 pixels of 6.1 microns square. While the HIRISE MEGA is a three megapixel chip with 2160x1440 pixels of 5.4 microns square. All three versions are filtered for colour and have two switchable pixel sensitivities, giving high quality or high sensitivity.

Returning to the concept of three colours being detected within each pixel, another company has also announced an image sensor with that facility — **Foveon**. This modified CMOS device is used in the newly announced Sigma SD9 digital SLR camera. The same principle is applied, of the colours being detected at different levels in the silicon. Foveon has also taken advantage of the concept by having what they call ‘Variable Pixel Size’, using all the pixels for maximum resolution, or combining adjacent pixels for greater sensitivity or flexible video recording (it is both a still and a video sensor); this is not a new concept, it has been used under the name of ‘binning’ for many years, but the necessity for maintaining the correct colour relationship between filtered pixels has until now limited its use to specialised applications. The F7 sensor used in the Sigma camera measures 20.7x13.8mm, and has a total of 2304x1536 pixels with a 9.12 micron pitch. For video, frame rates can vary from 2 per second at full resolution to 25 at 576x384. A second sensor, the F10, is expected to be launched later this year. This will be 6.4x4.8mm, with a 5 micron pixel pitch, and a total of 1344 x 1024 pixels. With a frame rate of 10 for full array up to 30 for VGA resolution. Unlike the F7 it will also have three integrated 12-bit A/D converters running at up to 20MHz; with red, green and blue multiplexed onto a single 10-bit tri-state output data bus, in addition to the RGB analog outputs. The Foveon sensor is shown in the **introductory photograph**.

FillFactory and TI

CMOS active pixels with their internal transistors can also have their sensitivity improved. FillFactory have developed what they call a **High Fill Factor N-Well Pixel**. In addition to the fill factor, the quantum efficiency has also to be taken into account (this being the number of electrons that are generated in the pixel for every impinging photon, relative to its spectral response). Photons can be lost by reflection on dielectrics, and by not being absorbed; while those electrons that are generated may not be collected and integrated, or the electron-hole pairs may recombine. As pixels are being made smaller, in order to increase the resolution, the problem is a growing one.

What FillFactory have done is to allow the photodiode to detect the generated electrons in the epitaxial layer beneath the surface of the whole pixel (including under the insensitive areas), electrons that would normally be lost. These are then channelled by electrostatic barriers that shield them from the active pixel circuitry and substrate to the photodiode junction (see **Figure 6**). Apparently, virtually all the escaped electrons will diffuse down that drain; giving a fill factor of 100%, with no other changes to the pixel structure. The only disadvantage is greater pixel crosstalk because of a 16% chance of electrons diffusing into an adjacent pixel (although this can be counteracted).

CCDs have not been left out when it comes to increasing sensitivity, either. Texas Instruments have developed a frame transfer sensor with integral image intensification which they call **Single Photon Detection** (SPD). In this the charge is multiplied before it is converted to an output voltage. The charge carrier multiplication (CCM) is achieved by creating a high-field region between two adjacent gates, and injecting the charge into this field. When electrons traverse the high-field region, they gain energy and when a certain threshold is exceeded impact ionisation occurs. This happens in the horizontal shift register as the charges are being read out. The register is twice the normal length (see **Figure 7**): the first half is of conventional design that interfaces with the storage area as in any other frame transfer CCD; the second half, however, includes 400 CCM stages. Here, new electron-hole pairs are created by impact ionisation; the new electrons are collected in potential wells where they are added to the original charge, and the holes are drained off to the substrate. This process is repeated for each stage until the greatly increased charge reaches the output amplifier.

The first SPD CCD has just been launched; and has a dynamic range of more than 90 dB. According to Texas Instruments, the SPD technology can be applied to any known image sensor architecture to increase the sensitivity of standard CCD imagers by more than 50 times. Additionally, it extends the spectral range so that it can cover from 200 nm (deep UV) to 1,100 nm (near IR).

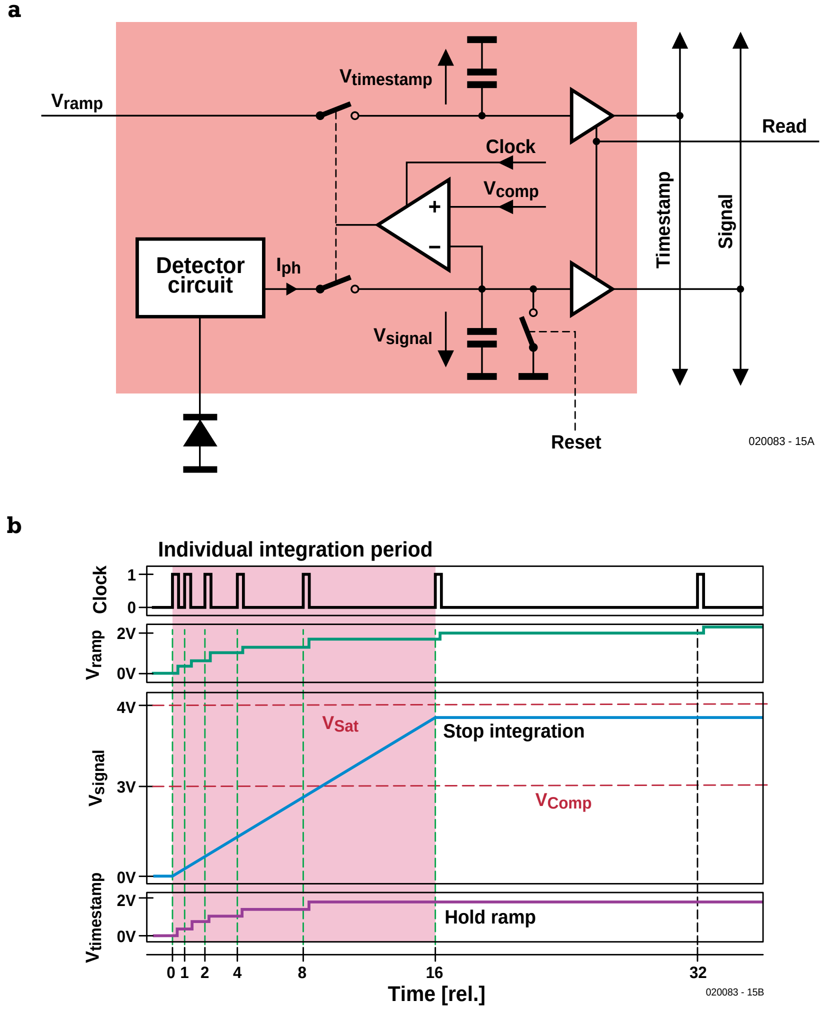


Figure 5. LARS pixel (5a) and its timing arrangement (5b) (courtesy Silicon Vision).

Cyborg vision?

Finally, there is a class of image sensor that borrows from the vertebrate eye — the **foveal** sensor. This type is being increasingly used for machine vision, such as robots. Instead of the pixels being evenly spread across the surface, there is a central fovea with smaller pixels in which the resolution is greater. In operation, the

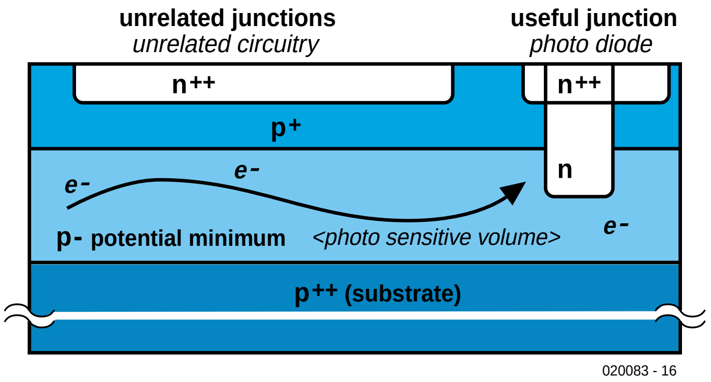


Figure 6. Electron crowding method to improve the fill factor (courtesy Fillfactory).

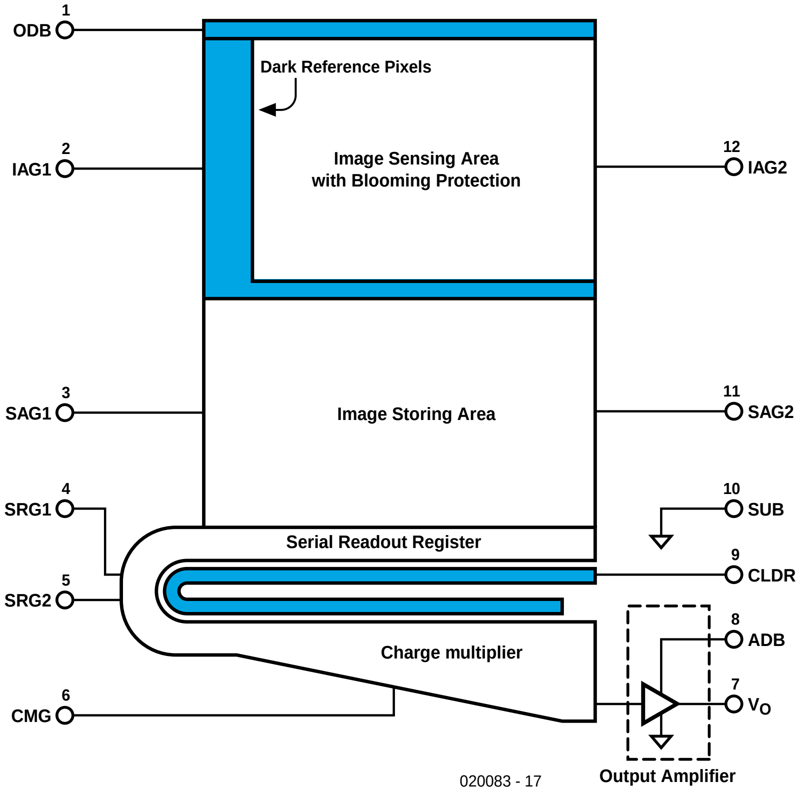


Figure 7. Single Photon Detection (SPD) is expected to make it to commercial CCDs soon (courtesy Texas Instruments).

peripheral area finds something of interest, and then directs the fovea to get a detailed look at it. The field of view can be broad; but because only the fovea provides the details the total number of pixels is much reduced by comparison with a conventional arrangement — as is the focal plane processing, where used. Many designs are circular, with the pixels arranged radially, although square ones have been produced, too. Most of them are CMOS designs: not only to easily incorporate on-chip processing, but also because prototypes can be produced at low cost. The latter being an obvious advantage, particularly as many of the designs are being produced in universities for research purposes.

Conclusion

Solid state image sensors have come a long way since the first cameras in 1972. The initial progress was painfully slow; but now there are genuine alternatives to the CCD and more recent CMOS, and generally the quality is rising in inverse proportion to the cost. And the new ‘smart’ sensors can perform functions that would formerly have required a lot of off-chip processing to achieve.

(020083-1)